yang_seonhyeHW27

```
library(data.table)
library(MASS)
data1 <- fread("http://users.stat.ufl.edu/~winner/data/muscle2.dat")
muscle <- data.table("muscle2.dat")
colnames(data1)[colnames(data1)=="V1"] <- "BodyMass"
colnames(data1)[colnames(data1)=="V2"] <- "WorkLevel"
colnames(data1)[colnames(data1)=="V3"] <- "HeatOutput"
```

```
attach(data1, warn.conflicts = F)
fit <- lm(HeatOutput~BodyMass+WorkLevel)
summary(fit)
```

```
##
## Call:
## lm(formula = HeatOutput ~ BodyMass + WorkLevel)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -282.0 -109.2    9.1  123.9  235.9
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  977.425    376.053   2.599 0.013723 *
## BodyMass      17.778      4.943   3.597 0.001011 **
## WorkLevel      6.244      1.522   4.102 0.000242 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 147.1 on 34 degrees of freedom
## Multiple R-squared:  0.4922, Adjusted R-squared:  0.4624
## F-statistic: 16.48 on 2 and 34 DF,  p-value: 9.914e-06
```

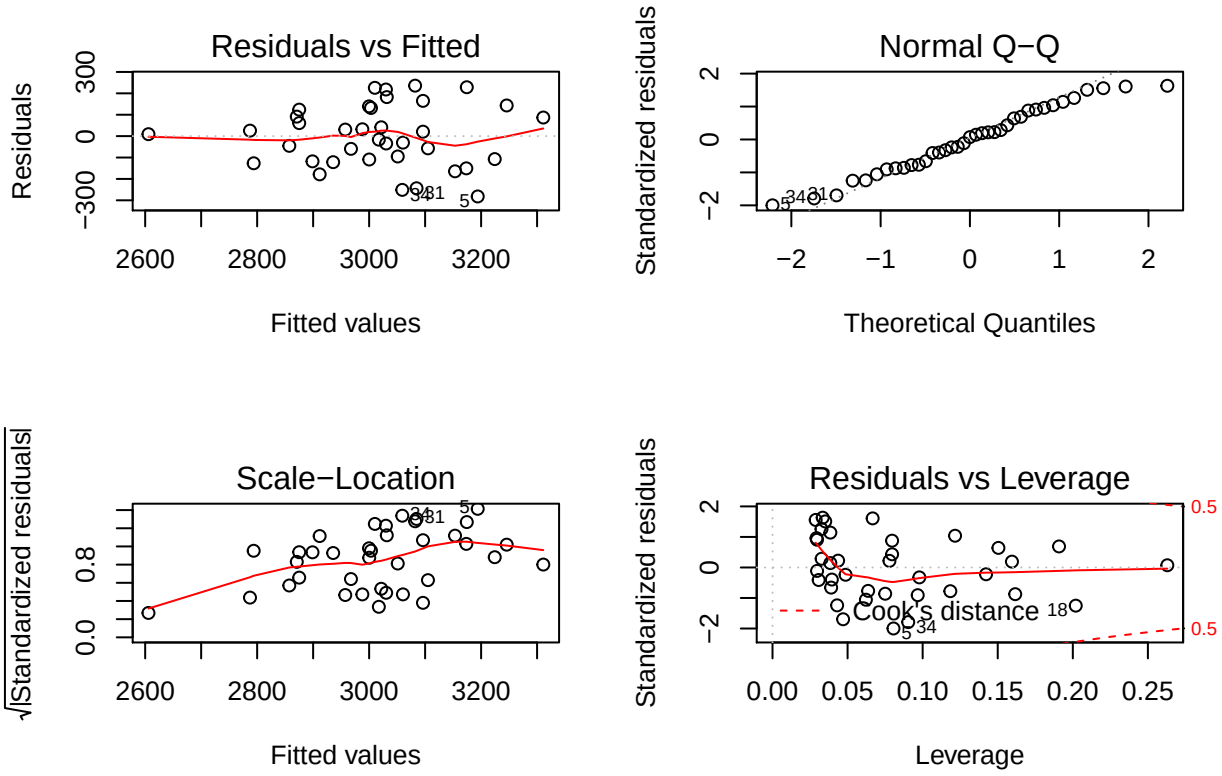
```
critval = qt(0.05/(2*nobs(fit)), df=df.residual(fit)-1, lower=FALSE)
which(abs(rstudent(fit)) > critval)
```

```
## named integer(0)
```

```
#there's no outliers
```

looking at our p-values, both predictors have a value less than 0.05 indicating they are significant to predicting heat output.

```
par(mfrow = c(2, 2))
plot(fit)
```



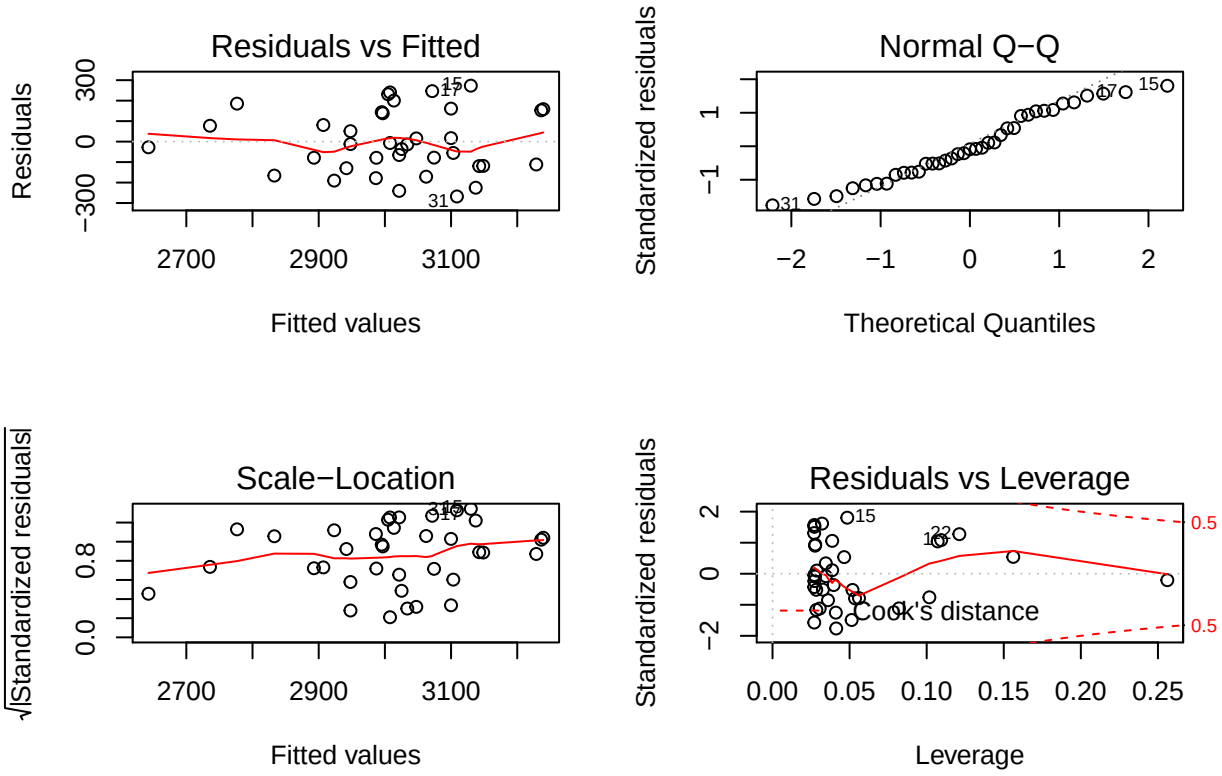

```
shapiro.test(fit$resid)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fit$resid
## W = 0.97134, p-value = 0.446
```

However, looking at our dianostic plot, we can see that linearlity is violated. Looking at the normal Q-Q, normailty is not violated. Looking at scale-location plot, residuals are not equally spread out along the ranges of predictors. Looking at Residuals vs Leverage, there doesn't seem to be any outliers.

We have a p-value greater than 0.05 and this means we cannot reject the null. Therefore, there's evidence for normality.

```
fit1 <- lm(HeatOutput ~ log(BodyMass+WorkLevel)) #final model
par(mfrow = c(2, 2))
plot(fit1)
```




```
#Looking at Scale-location, constant variance is not violated.
summary(fit1)
```

```
##
## Call:
## lm(formula = HeatOutput ~ log(BodyMass + WorkLevel))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -267.83 -118.57  -13.87  143.74  273.16
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 -4688.3     1534.2  -3.056  0.00428 **
## log(BodyMass + WorkLevel)    1454.3      289.6   5.022  1.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 155.1 on 35 degrees of freedom
## Multiple R-squared:  0.4188, Adjusted R-squared:  0.4022
## F-statistic: 25.22 on 1 and 35 DF,  p-value: 1.5e-05
```

```
critval = qt(0.05/(2*nobs(fit1)), df=df.residual(fit1)-1, lower=FALSE)
which(abs(rstudent(fit1)) > critval)
```

```
## named integer(0)
```

```
shapiro.test(fit1$resid)
```

```
##
##  Shapiro-Wilk normality test
```

```
##
## data:  fit1$resid
## W = 0.96022, p-value = 0.2052
```

```
#there are no outliers
```

We have a p-value greater than 0.05 and this means we cannot reject the null. Therefore, there's evidence for normality.